

## "Die Methode liegt im Code": Routinen und digitale Methoden im Datenjournalismus

Ausserhofer, Julian

Erstveröffentlichung / Primary Publication  
Sammelwerksbeitrag / collection article

### Empfohlene Zitierung / Suggested Citation:

Ausserhofer, J. (2015). "Die Methode liegt im Code": Routinen und digitale Methoden im Datenjournalismus. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 87-111). Berlin <https://doi.org/10.17174/dcr.v2.5>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:  
<https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:  
<https://creativecommons.org/licenses/by/4.0>

**Empfohlene Zitierung:** Ausserhofer, J. (2015). „Die Methode liegt im Code“: Routinen und digitale Methoden im Datenjournalismus. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 87-111). doi: 10.17174/dcr.v2.5

**Zusammenfassung:** Digitale Methoden haben in den letzten Jahren nicht nur in die Sozial- und Geisteswissenschaften Einzug gehalten. Auch der Journalismus hat seine Verfahren digitalisiert und seine Recherchepraktiken sowie seine Darstellungsformen entsprechend angepasst. Immer öfter bedienen sich Journalist:Innen für ihre Berichterstattung des Werkzeugkastens der digitalen Methoden. Am deutlichsten zeigt sich das am Datenjournalismus, der seit einem Jahrzehnt einen enormen Aufschwung erfahren hat. Dieser Beitrag beleuchtet digitale Methoden im Datenjournalismus. Zu diesem Zweck werden hier zunächst aus den Science and Technology Studies und der Digital-Methods-Forschung heraus die Grundzüge einer datenjournalistischen Methodologie entwickelt. Auf Grundlage von Redaktionsbesuchen und elf semistrukturierten Interviews mit Praktiker:Innen werden im Anschluss konkrete Arbeitsabläufe und Akteure in datenjournalistischen Methoden beschrieben. Es wird erläutert, wie datenjournalistische Kollaborationen im Newsroom funktionieren, wie sich dort Routinen etablieren und welche Technologien dabei vermitteln. Abschließend werden Empfehlungen zur Weiterentwicklung digitaler Methoden in der Kommunikationswissenschaft und im Datenjournalismus formuliert.

**Lizenz:** Creative Commons Attribution 4.0 (CC-BY 4.0)

*Julian Ausserhofer*

## „Die Methode liegt im Code“

### Routinen und digitale Methoden im Datenjournalismus

#### 1 Einleitung

Am 15. April 2013 explodierten zwei Bomben in der Zielgeraden des Marathons von Boston. Drei Menschen verloren ihr Leben, mehr als 260 weitere wurden teils schwer verletzt. Während vor Ort Sicherheits- und Rettungsmaßnahmen eingeleitet wurden, machte sich die Polizei an die Aufklärung der Tat. Auch ReporterInnen und Laien versuchten zur Aufklärung beizutragen und sammelten auf Online-Plattformen sämtliche verfügbaren Informationen zum Attentat. Sie analysierten Fotos und Videos, transkribierten den Polizeifunk, glichen die Ergebnisse mit Datenbanken, Karten und Social Networks ab. Nachdem die Polizei die Namen von zwei Verdächtigen veröffentlicht hatte, dauerte es nicht lange bis auch deren Social-Media-Profile öffentlich bekannt waren (Madrigal, 2013). Viele Nachrichtenorganisationen zitierten in der Folge die Tweets von Dzhokhar Tsarnaev, dem jüngeren der zwei Brüder (z.B. Preston & Roston, 2013). Die Medienhäuser behandelten Twitter als eine weitere Textquelle neben vielen anderen. Wenige andere Nachrichtenseiten wählten einen alternativen Ansatz: Um mehr über den Verdächtigen und dessen Lebensumstände zu erfahren, beschäftigten sie sich mit den Metadaten seines Twitter-Profiles. So extrapolierte die Website *Quartz* etwa die Zeitpunkte, zu denen der mutmaßliche Attentäter im Laufe von

eineinhalb Jahren seine rund tausend Tweets abgesendet hatte (Seward et al., 2013). Das Resultat: Eine Aufschlüsselung nach Wochentagen und Stunden über Tsarnaevs durchschnittliche Twitter-Aktivität, woraus die Autoren ableiteten, wann er für gewöhnlich schlafen würde.<sup>1</sup> Die Nachrichtenseite *Digg* veröffentlichte wenige Tage später eine umfassende Analyse von Tsarnaevs Twitternetzwerk und des *#FreeJahar*-Hashtags, unter dem die Freilassung und Schuld von Tsarnaev diskutiert wurden (Lotan, 2013). Mit Hilfe einer Clusteranalyse wurden Gruppen identifiziert, mit denen Tsarnaev auf Twitter verbunden war. Ausgewählte anonymisierte Zitate aus deren Timelines komplettierten das Bild über einen Informationskanal, den der mutmaßliche Attentäter wohl stark verwendet hat.

Derartige „computerunterstützte Explorationen im Journalismus“ (Gynnild, 2014, S. 3) haben in den vergangenen Jahren stark zugenommen. Der experimentelle Einsatz von Algorithmen, Daten und sozialwissenschaftlichen Methoden in der Nachrichtenproduktion – von der Recherche über die Datenanalyse bis hin zur Visualisierung – wird im deutschsprachigen Raum vor allem unter dem Begriff Datenjournalismus diskutiert (für eine Begriffstypologie siehe Coddington, 2014).<sup>2</sup> Wie die eingangs beschriebenen Twitter-Analysen beispielhaft demonstrieren, werden im Datenjournalismus immer wieder Daten verwendet, die von Webplattformen generiert werden. Auch wenn es nicht so titulierte wird, werden auch „digitale Methoden“ eingesetzt: DatenjournalistInnen untersuchen wiederkehrend Themen, Akteure oder Strukturen von Online-Kommunikation. In leitenden Fragestellungen wie auch in eingesetzten Methoden zeigen sich starke

- 1 Der Vollständigkeit halber soll nicht unerwähnt bleiben, dass die Autoren später im Artikel derartige Recherchen durchaus kritisch hinterfragen: „Where it was once only reporters and the police who dug up information about people of interest, a whole nation is at it today. [...] But those caveats are almost beside the point. We don't know what we think we know because these digital details don't connect the dots; they merely draw the dots. They offer trivia but not insight.“ (Seward et al., 2013)
- 2 Neben der Bezeichnung „Datenjournalismus“ haben sich je nach Medium, Region oder disziplinärem Hintergrund auch andere Begriffe etabliert: Data-driven journalism, precision journalism, computational journalism, database journalism, structured journalism und computational journalism sind nur einige der vielen Ausdrücke, die im Umlauf sind. Mit graduellen Unterschieden benennen sie die gleichen Tätigkeiten. Aus Gründen der Kürze und Einfachheit verwende ich im Folgenden auch immer wieder die Bezeichnung „DatenjournalistIn“, wohlwissend dass dieser Begriff so wie auch Datenjournalismus Unschärfen beinhaltet und es „den“ Datenjournalisten nicht gibt.

Parallelen zur sozialwissenschaftlichen Internetforschung. Aber ähnlich wie beim „Geschwisterpaar“ empirischer Journalismus und empirische Sozialwissenschaft (Haas, 1999) gibt es auch deutliche Unterschiede zwischen Datenjournalismus und sozialwissenschaftlicher Internetforschung. Das wechselseitige Verständnis des jeweils anderen Metiers kann indes fruchtbare Impulse zur Verbesserung des eigenen Metiers liefern (Rühl, 1981, zitiert nach Haas, 1999, S. 111).

Vor diesem Hintergrund beschreibt dieser Aufsatz Routinen, Arbeitsabläufe und Technologien im Datenjournalismus. Für die Kommunikationswissenschaft will dieser Text ein „Blick über den Tellerrand“ sein: Eine Analyse der digitalen Methoden des Datenjournalismus soll der sozialwissenschaftlichen Internetforschung alternative Herangehensweisen zeigen. Für Datenjournalismus-Interessierte bietet der Beitrag eine strukturierte Beschreibung der Arbeitsabläufe in datenjournalistischen Projekten (Abschnitt 5). Bevor das passiert, soll jedoch noch erörtert werden, was unter „digitalen Methoden“ verstanden wird und wie sich diese erforschen lassen (Abschnitt 2). Auch wird dargelegt, was die Forschung bislang zu datenjournalistischen Routinen und Methoden weiß (Abschnitt 3) und wie für diesen Forschungsbeitrag vorgegangen wurde (Abschnitt 4).

## **2 Methodologischer Zugriff: Digitale Methoden, ihre Verteilung und Erforschung**

Wenn im Folgenden von digitalen Methoden die Rede ist, dann bezieht sich das auf den Entwurf von Richard Rogers (2009, 2010, 2013) und die von ihm gegründete *Digital Methods Initiative* (DMI) an der Universität Amsterdam. Formuliert als Methodologie für die Sozial- und Geisteswissenschaften unterscheidet Rogers zwischen digitalisierten und digitalen Methoden: Erstere umfassen ursprünglich analoge Forschungsansätze, die für das Internet adaptiert wurden – Ethnografien oder Fragebögen etwa. Zumeist beschäftigen sich Forschungsprojekte, die diese Methoden zur Grundlage haben, mit netzkulturellen Phänomen oder sie nutzen das Web schlicht als Kommunikationsmedium im Forschungsprozess. Digitale Methoden hingegen sind Teil des „Computational Turn“ (Berry, 2011), sie sind „nativ digital“ (R. Rogers, 2014), sie wurden also speziell für Online-Forschungsumgebungen entwickelt und haben kein direktes analoges Vorbild. Auch die Forschungsperspektive ist bei digitalen Methoden eine andere: Das Web wird

nicht als virtueller Raum gesehen, dessen Kultur durch Forschung durchdrungen wird; die Analyse von Online-Umgebungen soll vielmehr dazu beitragen, Aspekte gesamtgesellschaftlicher Kommunikation und Kultur zu verstehen: „Rather, in studying the online, we make and ground findings about society and culture with the Internet. Thus, the Internet is a research site where one can ground findings about reality.” (R. Rogers, 2010, S. 243)

So beschäftigen sich Digital-Methods-Forschungsprojekte etwa mit der Rolle von Bots<sup>3</sup> in der Wissensakkumulation bei Wikipedia (Niederer & van Dijck, 2010), den Plattformen der holländischen Blogosphäre (Weltevrede & Helmond, 2012) oder den Online-Kontroversen zum Klimawandel (Niederer, 2013). Die Forschungsinteressen sind durch einen starken Fokus auf Plattformen und deren digitale Objekte wie Hyperlinks, Likes oder Tags gekennzeichnet. Automatisierungen wie etwa das „Crawlen“ von Netzwerken und Websites oder das „Scrapen“<sup>4</sup> von Social-Media-Inhalten sind typische Datenbeschaffungsvorgänge. Die Ergebnispräsentationen sind geprägt von Visualisierungen wie Netzwerken, Wortwolken oder Zeitstrahlen.

Unter welchen methodologischen Prämissen lassen sich digitale Methoden erforschen? Nortje Marres (2012), die starke Verbindungen zur *Digital Methods Initiative* hat, argumentiert unter Rekurs auf unter anderem Bruno Latour (1988) und John Law (2004), dass Methoden ein verteiltes Unterfangen darstellen. Anschließend an die Tradition der Akteur-Netzwerk-Theorie (ANT) sieht sie Methoden als gemeinsame Entwicklung von ForscherInnen, Forschungssubjekten und -objekten sowie Technologien. Neues Wissen und Innovation in Forschungsprozessen werden durch komplexe Aushandlungsprozesse arbeitsteilig zwischen ebendiesen Akteuren entwickelt. Die Digitalisierung, so Marres, habe sozialwissenschaftliche Methoden stark verändert und etwa Online-Plattformen als wichtige Akteure mit Forschungskapazität etabliert. Weitere Agenten digitaler Methoden seien neben digitalen Objekten (R. Rogers, 2009) unter anderem auch Datenbanken, Designagenturen, Algorithmen, IT-Unternehmen oder BürgerwissenschaftlerInnen.

3 Ein Bot ist eine Softwareanwendung, die automatische einfache, repetitive Aufgaben im Internet ausführt.

4 Crawler und Scraper sind Skripte, die ganze Webseiten bzw. Datenbanken oder Teile davon abgreifen, auf eine lokale Repräsentanz kopieren und für die weitere Analyse aufbereiten.

Führt man die von Marres aufgebrachte Methodologie konsequent weiter, dann stellt sich vor dem Hintergrund dieses Aufsatzes die Frage, inwieweit man die Handlungsmacht von Agenten, die an digitalen Methoden im Datenjournalismus beteiligt sind, konsequent operationalisieren und damit erforschen kann. Es liegt nahe, auf die Aushandlungsprozesse und die daraus resultierende verbundene Arbeitsteilung zwischen den einzelnen Akteuren zu fokussieren. Um die Handlungsmacht der involvierten Plattformen und Software-Agenten zu bestimmen, eignen sich Forschungsansätze wie jener der Software Studies (Manovich, 2013) oder der digitalen Methoden selbst – unter der Voraussetzung, dass eine selbst-reflexive Offenlegung<sup>5</sup> des Vorgehens stattfindet. Zur Analyse der physischen Situationen in den Redaktionen bieten sich neben Newsroom-Ethnografien (z.B. Domingo & Paterson, 2011) Forschungsprogramme wie die ANT (Latour, 2005) oder die von Robert Schmidt (2012) entwickelte Soziologie der Praktiken an, die verschiedene Stränge sozialwissenschaftlicher Forschung inkorporiert. Schmidt vereint für seinen praxeologischen Entwurf Ansätze der Ethnomethodologie, der Interaktionsanalyse, der ANT sowie Bourdieus ethnologischen Feldstudien. Im Mittelpunkt von Schmidts Ansatz stehen die Beobachtbarkeit sozialer Praktiken sowie die Körperlichkeit und Materialität bei Interaktionen.

Im Ergebnisteil versucht diese Arbeit, diese Perspektiven auf digitale Methoden und die Möglichkeit ihrer Erforschung in die Praxis umzusetzen und insbesondere die Bandbreite und die Eigenschaften der Akteure, die an datenjournalistischen Methoden beteiligt sind, zu ergründen. Empirisch wird neben Beobachtungsprotokollen auf Interviewmaterial zurückgegriffen.

### **3 Forschungsstand: Datenjournalistische Routinen**

Die Dichotomie von Euphorie und Skepsis, die regelmäßig in der Frühphase von gesellschaftlichen Debatten über neue technologiegetriebene Praktiken auftritt (Münch & Schmidt, 2005), ist auch beim Datenjournalismus festzustellen;

5 Die selbstreflexive Offenlegung umfasst die Transparenzmachung von Entscheidungen und Interpretationen im Forschungsprozess sowie die Veröffentlichung von Forschungsartefakten wie Bibliografien, Notizen, Algorithmen und rohen sowie prozessierten Daten (siehe dazu Ausserhofer, 2015; Rieder & Röhle, 2012).

wobei in der (halb-)akademischen Literatur die optimistische Sichtweise überwiegt (Anderson, 2013) – ganz ähnlich wie bei den Diskursen zu partizipativem Journalismus (Borger et al., 2013). Frühe Veröffentlichungen zu Datenjournalismus weisen insbesondere auf die sich durch die neue Technologie eröffnenden Potenziale für den Journalismus und die Demokratie im Allgemeinen hin (exemplarisch: Hamilton & Turner, 2009; Cohen et al. 2011; Flew et al., 2010).

Eine ähnliche Perspektive zeigte bereits Philip Meyer (2002/1973), der aufgrund seiner journalistischen Arbeiten und des Referenzhandbuchs „Precision Journalism“ als einer der Pioniere von Datenjournalismus gilt. Er beschreibt, wie JournalistInnen Statistik und andere quantitative sozialwissenschaftliche Verfahren für die Berichterstattung verwenden können. Darüber hinaus liefert er erstmals ein Workflow-Modell, das den Ablauf datengetriebener journalistischer Projekte operationalisiert. Auch andere haben derartige Schemata entwickelt (z.B. Bradshaw, 2011a, 2011b; S. Rogers, 2011; Lorenz, 2010; Weinacht & Spiller, 2014), in denen sie Methoden und Prozesse dokumentieren und visualisieren. Diese Schemata sind unterschiedlich detailliert und haben auch verschiedene Schwerpunkte, erklären im Kern jedoch folgende Prozesse: a) Datenbeschaffung und -bereinigung, b) Storytelling, Entwicklung und Visualisierung, c) Veröffentlichung und Distribution.

Mit den Routinen und Methoden in Nachrichtenorganisationen (neben anderen Aspekten) beschäftigen sich auch einige jüngere, vor allem auf qualitativer Forschung aufbauende Veröffentlichungen. Es sind meist Fallstudien zu einzelnen Redaktionen oder Projekten (Royal, 2012; Parasie & Dagiral, 2013; Weber & Rall, 2013; Young & Hermida, 2014), oder Analysen der Situation von Datenjournalismus in einem Land oder einer Sprachregion (Weber & Rall, 2013; Applgren & Nygren, 2014; De Maeyer et al., 2014; Fink & Anderson, 2014; Karlsen & Stavelin, 2014; Weinacht & Spiller, 2014). In den Arbeiten zeigt sich, dass zwar so manche Unterschiede in den analysierten Häusern und Ländern (USA, Belgien, Schweden, Norwegen, Deutschland) bestehen – Stichwort Hierarchie und Organisationskultur –, im Hinblick auf Epistemologie, Praktiken und technologische Entwicklungsstandards aber durchaus viele Ähnlichkeiten<sup>6</sup> bestehen: So existieren

6 Die Homogenität in den Ergebnissen in den bisherigen Forschung hängt vermutlich auch mit der Erhebungsmethode „Interview“ zusammen, in der es die Tendenz zu erwünschten Antworten gibt, die wiederum vom normativen transnationalen Diskurs über Datenjournalismus beeinflusst sind (Gynnild, 2014; De Maeyer et al., 2014).



einerseits einige wenige datenjournalistische Leuchtturmprojekte und -redaktionen, die auch häufig im Blickpunkt der Analysen stehen, und andererseits viele (Individuen in) Organisationen, die mit Ressourcenknappheit kämpfen.

Diese Dichotomie wird auch in der Art der realisierten Projekte offenbar – zum einen gibt es jenen Datenjournalismus, der von einer Einzelperson im „Tagesgeschäft“ produziert wird, zum anderen Investigativprojekte, mit denen Teams in Arbeitsteilung oft wochen- oder monatelang beschäftigt sind (Fink & Anderson, 2014). Ähnlich wie im „traditionellen“ Investigativjournalismus bilden sich dabei die interdisziplinären Teams je nach Projektanforderung neu, oft auch unter Einbeziehung von externen Dienstleistern. Deadlines treten in den Hintergrund, im Vordergrund stehen die Vollständigkeit des Beitrags und der Wunsch, mit jedem Projekt das technologische Potenzial etwas weiter auszuloten (Royal, 2012). Dementsprechend gelten viele DatenjournalistInnen in ihren Redaktionen als technologische Avantgarde, als „Nerds“: Sie zeigen Affinität zu Transparenz-Ideologien aus der IT-Kultur und versuchen Methodologien wie Hacking in die Redaktionen zu tragen (Parasie & Dagiral, 2013; Howard, 2014; Weinacht & Spiller, 2014). Diese Arbeitskulturen sind mit jenen im „traditionellen“ Journalismus nicht völlig kompatibel, weshalb sich Redaktionen durchaus schwer damit tun, Routinen zu etablieren (Karlsen & Stavelin, 2014).

#### **4      Forschungsfokus und -design**

Ausgehend von den obigen Ausführungen rund um digitale Methoden und deren Erforschung sowie um datenjournalistische Routinen sollen hier das Forschungsinteresse und -design spezifiziert werden. In Fortführung der methodologischen Argumentation und Literaturdiskussion ergeben sich folgende Forschungsfragen, die in dieser Reihenfolge im Ergebnisteil beantwortet werden:

- *FF1: Wie funktioniert Zusammenarbeit bei datenjournalistischen Projekten?*
- *FF2: Wie sehen datenjournalistische Routinen und Arbeitsabläufe aus?*
- *FF3: Welche Akteure (im Sinne der Akteur-Netzwerk-Theorie) beteiligen sich an digitalen datenjournalistischen Methoden?*
- *FF4: Wie werden digitale datenjournalistische Methoden stabilisiert?*

Um diese Forschungsfragen zu beantworten, wurden zwischen Ende August 2013 und Mitte Mai 2014 semi-strukturierte Interviews mit elf Personen durchgeführt, die überwiegend einer Tätigkeit nachgingen, die mit Datenjournalismus assoziiert wird.<sup>7</sup> Diese übten sie entweder in Redaktionen (acht der elf Befragten) oder für spezialisierte Datenjournalismus-Agenturen (3) aus – bei insgesamt sieben unterschiedlichen Organisationen.<sup>8</sup> Sie wurden um ein Interview gebeten, weil sie in der Berichterstattung über Datenjournalismus immer wieder genannt worden waren oder weil sie sich auf Twitter unter dem einschlägigen Hashtag „#ddj“ geäußert hatten und sich in ihrem dortigen Profil entsprechend darstellten. Zehn der Interviewten arbeiteten im deutschsprachigen Raum, eine Person in Großbritannien.<sup>9</sup>

Die Interviews wurden nach einem schwach strukturierten Leitfaden face-to-face in den Redaktionen bzw. den jeweiligen Arbeitsumgebungen aufgezeichnet.<sup>10</sup> Zusätzlich zu den Interviews wurden nach den Besuchen Beobachtungsprotokolle verfasst, die Aspekte wie räumliche Ordnung, Gesten und Materialien in interpersonellen Interaktionen dokumentierten. Die durchschnittliche Gesprächsdauer betrug eine Stunde und zwei Minuten. Zu Beginn wurden die Interviewten nach ihrer Ausbildung, ihrem beruflichen Werdegang und ihrem Einstieg in den Datenjournalismus gefragt. Anschließend rückten die Akteure, Abläufe und Methoden bei datenjournalistischen Projekten in den Mittelpunkt des Gesprächs. Der gesamte Prozess, von Themenfindung über Produktionsbedingungen bis hin zur Dissemination datenjournalistischer Projekte wurde im Detail durchgegangen. Den Abschluss bildeten Reflexionen über digitale Methoden, die eigene

7 Die Befragten waren (in alphabetischer Reihenfolge): Mona Chalabi, Christina Elmer, Florian Gossy, Markus Hametner, Thomas Jöchler, Nicolaus Kayser-Bril, Mirko Lorenz, Lorenz Matzat, Maximilian Schäfer, Julius Tröger und Sascha Venohr.

8 Zum Zeitpunkt der Interviews arbeiteten die Personen bei den bzw. für die folgenden Organisationen (in alphabetischer Reihenfolge): Berliner Morgenpost, The Guardian, Journalism++, Open Data City, Der Spiegel, Der Standard, Die Zeit.

9 Wie groß die Dynamik in der Branche ist, zeigt sich am Umstand, dass mit Herbst 2014, zum Zeitpunkt der Niederschrift dieses Textes, wenige Wochen nach dem letzten Interview, mindestens sechs der Interviewten zwar noch in derselben Branche aber bereits in anderen Abteilungen oder bei anderen Organisationen tätig sind.

10 Ein Interview wurde per Skype geführt, zwei Personen wurden an anderen Orten interviewt.

Praxis und die allgemeine Situation des Datenjournalismus. Insbesondere wurde dabei auf die Nähe zur Wissenschaft, auf die Objektivität und auf die Ausbildungssituation von DatenjournalistInnen eingegangen.

Die Interviews wurden transkribiert und anschließend mit den Beobachtungsprotokollen in ein Programm zur computerunterstützten qualitativen Auswertung, den TAMS Analyzer (Weinstein, 2006), importiert. Die Analyse orientierte sich an den Konzepten der Grounded Theory (Berg, 2001; Glaser & Strauss, 1967; Strauss & Corbin, 1990) und wurde auf Basis der Transkripte entwickelt. Die Codes der Analyse wurden in einem iterativen Prozess auf die Gesamtheit des Materials angewendet.

## 5 Ergebnisse

*„Wie haltet ihr eure Methoden für später fest?*

*„Was meinst du damit?“*

*„Wie dokumentiert ihr etwa eure Visualisierung oder Datenakquise?“*

*„Das ist bei uns als Code vorhanden.“*

Aus einem Interview mit einem Datenjournalisten (I1)<sup>11</sup>

### 5.1 Interdisziplinäre Kollaborationen

Die Produktion eines datenjournalistischen Stücks ist in der Regel ein kollektives Unterfangen, erklärte die Mehrzahl der Interviewten. Viele unterschiedliche Akteure helfen mit, bis ein Projekt erscheint. „Den Datenjournalisten“ oder „die Datenjournalistin“ gibt es in den wenigsten Organisationen. Gewöhnlich werden datenjournalistische Stücke arbeitsteilig von SpezialistInnen produziert. Ihre Expertise liegt in Programmierung, Statistik, Software-Projektmanagement, interaktiver Grafik, nutzerzentriertem Design, Kartographie, Fact-Checking oder in einem spezifischen Themengebiet. Viele der Interviewten sind ExpertInnen in mehreren Domänen. Mit geringer externer Unterstützung können sie allein datenjournalistische Stücke veröffentlichen. Zentral dafür sind das Beherrschen

11 Die Zitate wurden der Schriftsprache angepasst, inhaltlich jedoch nicht verändert. Zwei der Interviews wurden auf Englisch geführt – die Zitate wurden übersetzt. Auch wurden die Namen der Interviewten durch Nummern ersetzt.

dreier Kompetenzen: Datenanalyse bzw. Programmierung, Design sowie Formulierungsstärke. Mängel in Visualisierungsfähigkeiten können in vielen Fällen durch den Einsatz webbasierter Tools kompensiert werden.

Die finanziellen Probleme, von denen viele Medienhäuser betroffen sind, beeinträchtigen auch die Personalpolitik im Datenjournalismus. So wie im Textjournalismus ist es im Datenjournalismus nichts Ungewöhnliches, für bestimmte Aufgaben externe Freischaffende zu beauftragen. Nur ressourcenstärkere Nachrichtenorganisationen leisten es sich, Vollzeitkräfte oder sogar Teams zu beschäftigen, die sich allein um datenjournalistische Projekte kümmern. Viele, die sich in Redaktionen an journalistischen Datenprojekten beteiligen, interessieren sich privat für Informationsvisualisierung oder Datenbanken. Sie sind ReporterInnen oder RedakteurInnen, häufig auch MitarbeiterInnen der Grafik-, Archiv- oder Entwicklungsabteilung, die sich innerhalb der Organisation ein neues Aufgabenfeld suchen und einen Teil ihrer Arbeitszeit auf einschlägige Projekte verwenden. Je nach Ressourcen und Projektanforderungen beteiligen sie sich in den interdisziplinären Teams, die an den Projekten arbeiten. Die Leitung in diesen Teams übernimmt meist eine Person mit journalistischem Profil, die die Arbeit dann zwischen sich selbst und den Verantwortlichen für Datenanalyse, Programmierung, Visualisierung und dem Fachressort, dem die Geschichte thematisch zugeordnet ist, verteilt.

Insofern nehmen Personen, die überwiegend einer mit Datenjournalismus assoziierten Tätigkeit nachgehen, in der Aufgabenverteilung in ihren Organisationen eine selten klar definierte Funktion ein: Einerseits sind sie selbstständige ReporterInnen und RedakteurInnen, die selbst recherchierte Projekte publizieren, andererseits unterstützen sie RedakteurInnen, die keine ausgeprägte Datenkompetenz haben, bei ihren Veröffentlichungen. Sie sehen sich sowohl als Teil der Redaktion als auch als Teil der Supportdienste. RedakteurInnen „suchen regelmäßig die fachliche Beratung [bei uns]“, erzählt ein Datenjournalist (17). Und weiter: „Es ist mittlerweile so, dass wir in der Redaktion etabliert sind, dass wir uns als geschätzten Partner sehen, der Arbeit nicht wegmoderiert. [...] [Die Fachredakteure] wissen: Wenn wir sagen, etwas geht nicht, dann geht es wirklich nicht.“ Zu dieser Rolle zwischen Redaktionsmitglied und Dienstleister passt auch die Tatsache, dass DatenjournalistInnen interne Weiterbildungen durchführen, in denen sie die technischen Kompetenzen von FachredakteurInnen verbessern und ihnen Tools vermitteln, mit Hilfe derer sie ohne externe Hilfe Datenvisualisierungen anfertigen können.

Interessierte treffen sich in internen Fortbildungen genauso wie auf einschlägigen, klassisch abgehaltenen Konferenzen. Darüber hinaus werden auch informelle Treffen mit Gleichgesinnten organisiert – sowohl hausintern als auch extern. Eine besondere Rolle spielen Formate wie die Veranstaltungen der internationalen Graswurzelbewegung *Hacks/Hackers*, bei der „Hacks“, also JournalistInnen, mit „Hackern“<sup>12</sup> zusammenkommen, um sich auszutauschen und zusammenzuarbeiten (Lewis & Usher, 2014). Ein Reporter (I8), der sich hauptsächlich mit Datenprojekten beschäftigt und auf einem solchen Treffen ein Projekt vorgestellt hat, berichtet von seinen Erfahrungen im Anschluss an sein Referat: „Dort kommen die entsprechenden ‚Nerdfragen‘, die du sonst nicht bekommst. Und ich erzähle dann genau, wie wir das gemacht haben, helfe dem einen oder anderen auch damit.“ Die gegenseitige Unterstützung ist im Datenjournalismus laut den Interviewten stark ausgeprägt – auch über Medienunternehmen hinweg. Selbst MitarbeiterInnen von Unternehmen, die miteinander in Konkurrenz stehen, würden sich gegenseitig helfen. Das gute Klima der Kollaboration betonen alle Befragten in ihren Aussagen. Sie unterstützen einander etwa bei technischen Problemen oder inhaltlichen Fragen. „In dieser Szene geht das sehr schnell. Man bringt eine Frage auf Twitter ein und dann antworten einem drei Koryphäen aus den USA oder von wo auch immer“, erklärt ein Interviewter (I7). Twitter wird übereinstimmend als ein zentraler Disseminations- und Feedbackkanal genannt. Die Beteiligten organisieren ihre Kommunikation unter anderem durch die Verwendung des Hashtags „#ddj“, das für „data-driven journalism“ steht. Auch kommunizieren viele Akteure miteinander per Telefon, Skype oder E-Mail. Praktiken wie etwa das Einbetten fremder Inhalte oder die Übernahme von Visualisierungsmethoden oder von Stories mit lokalisiertem Inhalt sind üblich.

## 5.2 Routinen und Arbeitsabläufe

Als große Herausforderung nannte ein Großteil der Interviewten die Einführung von Routinen in die Produktion der datenjournalistischen Stücke. Viele

12 Hacker in diesem Kontext bezeichnet nicht Personen, die Sicherheitslücken in Computersystemen ausnützen, sondern technologieaffine Personen, die mit ihren Programmierkenntnissen spielerisch gesellschaftliche Probleme lösen wollen.

der Projekte haben experimentellen Charakter – selbst in Organisationen, die bereits seit mehreren Jahren Datenjournalismus betreiben. Die Installation von Routinen wird nicht zuletzt durch die Heterogenität der Projekte erschwert. Diese zeigt sich an drei Aspekten: Erstens sind Datenjournalismus-Verantwortliche in vielen Organisationen sowohl für tagesaktuelle als auch für langfristig angelegte Projekte zuständig. Da diese Projekte je nach Realisierungshorizont unterschiedliche Herangehensweisen erfordern, ist es schwer, Routinen zu etablieren. Zweitens macht das Themenspektrum, das von einer einzelnen Person oder einem Team behandelt wird, nicht vor Ressortgrenzen Halt: Mit Datenjournalismus betraute Personen arbeiten gewöhnlich an Fragestellungen und heterogenen Daten aus allen Themenbereichen einer Redaktion – von Außenpolitik über Sport bis zu Chronik. Drittens setzen viele Projekte den Einsatz neuer Tools, Software oder Visualisierungsbibliotheken voraus. „Man hat keine Strategie für die nächsten Jahre [...], man kann online nur die nächsten zwei Wochen planen und ‚wurstet sich durch‘ [bringt sich behelfsmäßig durch]“, resümiert ein Journalist (18).

Auch wenn kaum datenjournalistische Routinen etabliert werden, so können trotzdem abgeschlossene *Praktiken* identifiziert werden, die miteinander in Beziehung stehen: *Initiiert* werden datenjournalistische Projekte durch selbstständige Ideen und Recherchen sowie durch redaktionsinterne oder externe Anregungen und Einflüsse. Mehrere der Interviewten nennen die Veröffentlichung neuer Daten durch statistische Dienste als zentrales Ausgangsmoment ihrer Projekte. Die Datenerfassung passiert in komplexen Projekten oftmals automatisiert mit Hilfe von Scrapern oder Crawlern. Auch werden Daten auf „Vorrat“ gesammelt, obwohl zum Zeitpunkt der Akquise noch nicht klar ist, inwieweit sich die Daten für eine Projektrealisierung eignen. Ein freischaffender Datenjournalist (11) erläutert den Hintergrund:

„Am besten sammelst du einmal alles ein, was es gibt. Du kannst später immer noch filtern. Der Scraper läuft dann, du lässt ihn, wenn man so will, ‚reifen‘. Nach einem Monat oder zwei guckst du in die Daten und schaust, ob du Muster findest.“

Das Finden von Mustern ist Teil einer „explorative[n] Phase, in der man erst einmal schaut: Was sind das überhaupt für Daten? In welchem Zustand sind sie? Was für ein Aufwand ist es, sich damit zu beschäftigen? Welche Form kann das

am Ende annehmen?“, erklärt ein anderer Interviewter (I3). Weitere Praktiken, die ebenso am Beginn eines Projekts stehen, sind die Erstreinigung der Daten, die Formulierung von Hypothesen und das Herausbilden eines Narrativs. Nicht jedes Projekt enthält zwangsläufig jede Praxis – und auch die Reihenfolge unterscheidet sich von mal zu mal. In einigen Projekten etwa verzichten die ProduzentInnen auf das Formulieren von Hypothesen oder das Reinigen der Daten.<sup>13</sup>

Die *Explorationsphase* ist geprägt durch Experimentieren und Sondieren. Zitat eines Datenjournalisten (I8): „Es passiert ständig: Wir haben irgendwelche Ideen, werfen Datensätze rein und dabei kommt gar nichts raus.“ „Es geht viel um herumspielen“, charakterisiert ein anderer (I10) diese Phase. Mit der Formulierung eines Narrativs geht gewöhnlicherweise die Entwicklung der Datenvisualisierung einher. Dabei werden Skizzen per Hand aufs Papier gebracht, bei komplexeren Projekten werden auch so genannte Wireframes, also am Computer gezeichnete Entwürfe, angefertigt. Diese sollen dabei helfen, die Sicht der EndanwenderInnen von Anfang an in die Konzeption zu integrieren. Sofern mehr als eine Person an einem Projekt arbeitet, ist in diesem Produktionsabschnitt die Abstimmung zwischen den einzelnen Teammitgliedern besonders intensiv. Allerdings gibt es kaum formalisierte Besprechungen, dafür aber starken informellen Austausch: Indem die Teammitglieder auf Abbildungen deuten, Skizzen anfertigen oder unter Zuhilfenahme von Gestik ihre Ideen artikulieren, aktualisieren sie wechselseitig ihre Vorstellungen über das Projekt. Unterstützend kommen in den Organisationen verschiedene physische und elektronische Projektkoordinationstools wie Pinnwände, die den Projektstatus abbilden, gemeinsame Chatkanäle in IRC oder Skype, Wikis, oder Mailinglisten zum Einsatz.

So werden datenjournalistische Projekte in Arbeitsteilung und intensiver gegenseitiger Abstimmung weiterentwickelt, bis sie kurz vor der Finalisierung stehen. Es folgen interne Tests, Feedbackschleifen mit den Ideen- bzw. den Auftraggebern. Ein wichtiges Element, das enorme Ressourcen benötigt, ist die Adaption der Visualisierungen für verschiedene Plattformen, Betriebssysteme und Browser. Gerade in Projekten, bei der NutzerInnen mit den Inhalten interagieren können, sie also die Fakten erst durch weitere Eingaben erfassen können, stel-

13 Der im Englischen gängige Ausdruck „Data-driven Journalism“ impliziert die Anwendung von theorieinduzierten und hypothesenformulierenden Methoden (siehe auch Mahrt in diesem Band).

len unterschiedliche Darstellungs- und Interaktionstechnologien die EntwicklerInnen vor große Herausforderungen. Schließlich spielt der Veröffentlichungszeitpunkt gerade bei aufwändigeren Projekten eine wichtige Rolle, geben die Interviewten übereinstimmend an. Neben der prominenten Platzierung auf der Startseite eines Mediums setzen sie im „Rollout“ intensiv auf die Distributionseffekte sozialer Medien.

In der *Evaluierungsphase* wird die Qualität von Unternehmungen von den PraktikerInnen nicht zuletzt danach beurteilt, inwieweit sie die Grenzen des technisch Machbaren ausloten. Mit anderen Worten: Die Qualität eines Projekts misst sich nicht allein am Inhalt bzw. dessen erfolgreicher Vermittlung an das Publikum, sondern auch am Arbeitsaufwand, am Komplexitätsgrad der eingesetzten Daten und Technologien sowie an der Weiterentwicklung von Visualisierungsansätzen.

Was beeinträchtigt die Güte eines Projekts? So gut wie alle Befragten nennen Ressourcenknappheit, insbesondere den Mangel an Zeit und Geld, an erster Stelle. Weiters fehlt es in sehr spezialisierten Projekten zuweilen an technischer Expertise. Exemplarisch hierfür die Aussage einer Redakteurin (12):

„Das ist so eine Zwickmühle im tagesaktuellen Online-Journalismus –, dass das, was wir datenjournalistisch machen, in einer perfekten Welt natürlich viel länger und noch reflektierter und optisch opulenter daherkommen könnte, wir uns das aber oft einfach nicht leisten können.“

Die Knappheit an Ressourcen führt unter anderem dazu, dass die Nachbereitung der Projekte in vielen Fällen nicht so passiert, wie es sich die PraktikerInnen wünschen würden. Insbesondere betrifft das die Veröffentlichung der Daten, der selbst entwickelten Programme und Werkstattberichte, die von allen Interviewten als wichtige Elemente am Ende des datenjournalistischen Workflows genannt werden. Damit soll die eigene Arbeit transparent und nachvollziehbar gemacht werden sowie eventuell gemachte Fehler nach der Veröffentlichung korrigiert werden. Auch wenn DatenjournalistInnen diese selbst gesteckten Ziele nicht immer erreichen, halten sie die Idee sehr hoch, möglichst viel Transparenz über das eigene Schaffen herzustellen. „Wir machen alles open source, weil wir sehen nicht, warum es anders sein sollte“, erklärt ein Gründer einer Datenjournalismussagentur (14) seine Motive.



### 5.3 Akteure digitaler datenjournalistischer Methoden

In den Beschreibungen der Arbeitsabläufe im vorhergehenden Abschnitt wurden bereits zahlreiche Akteure erwähnt, die an datenjournalistischen Methoden beteiligt sind. Die Mehrheit der Interviewten schildert, dass Software-Komponenten im gesamten Prozess zentrale Rollen einnehmen. In der Phase der Datenakquise sind es etwa so genannte „digitale Briefkästen“. Dabei handelt es sich um Webformulare, über die eine Organisation verschlüsselt und anonymisiert Daten empfangen kann. Auch Crawler und Scraper kommen zum Einsatz. Nicht zuletzt sind auch Schnittstellen zu Datenbanken zentrale Agenten der datenjournalistischen Methode, kontrollieren sie nicht nur den Zugang zu strukturierten Informationen, sondern formatieren gemeinsam mit den Datenbanken selbst und deren Nutzungsbedingungen auch die Weiterverarbeitung der Daten vor (Ausserhofer, 2015; Bodle, 2011; Bucher, 2013).

Weitere Hard- und Software-Agenten beeinflussen die lokale Datenaufbereitung und -visualisierung: In den Interviews wurden unter anderem Tabellenkalkulationsprogramme und Visualisierungswerkzeuge genannt. Die DatenjournalistInnen würden die eigenen Vorstellungen ständig anpassen, weil die Software-Anwendungen nur bestimmte Operationen zulassen. Durch Programmieren könne man diese Restriktionen jedoch weitestgehend überwinden. Ein Datenjournalist (I10) erläutert: „Die Software gibt vor, wie Geschichten auszusehen haben. Zugleich ist es ein großer Vorteil, dass wir im Newsroom einen Entwickler haben, der mit uns eigene Lösungen baut, damit wir uns nach den Daten richten können und nicht nach der Software.“

Daten werden auf lokalen Servern oder in der „Cloud“ abgelegt. Darüber hinaus dienen die Server auch als „Hosts“ für Datenscraper und -crawler sowie News-Apps. Äußerst wesentlich erscheinen Hosting-Dienste wie *GitHub*, mit Hilfe derer einerseits kollaborativ Programmcode entwickelt wird, andererseits Daten veröffentlicht werden. Immer wieder stören lokale Sicherheitsvorkehrungen und Kompatibilitätsprobleme die Arbeit an Projekten, zum Beispiel wenn der Laptop, auf dem die Visualisierung entwickelt wird, bestimmte Programme nicht ausführen kann oder wegen Sicherheitsbedenken von den lokalen Servern ausgesperrt wird – alltägliche Situationen, wie sie mehrfach geschildert wurden.

Eine nicht unwesentliche Rolle spielt auch die räumliche Position der DatenjournalistInnen in der Organisation. Aus den Aussagen der Interviewten und den

Beobachtungen während der Redaktionsbesuche lässt sich keine eindeutige Tendenz ablesen – auch weil die räumliche Situation in jeder Organisation unterschiedlich ist. Manche Akteure waren mitten im Newsroom stationiert, andere in der Entwicklungsabteilung, dritte waren ohne zugeteilten Arbeitsplatz, saßen mit ihrem Laptop an freien Tischen oder besetzten eigens für Recherche vorgesehene Räume. In den Agenturen dominierten Gemeinschaftsbüros.

Neben der physischen Position in der Organisation während des Entwicklungsprozesses ist auch der Platz der Veröffentlichung fertiger Projekte auf der Website von großem Belang für die Interviewten. Dies wird gewöhnlich von leitenden Angestellten entschieden. Eine prominente Platzierung auf der Start- oder Ressortseite, zu Zeiten, in denen die Seite oft aufgerufen wird, interpretieren DatenjournalistInnen als Wertschätzung der Redaktionsleitung gegenüber der geleisteten Arbeit. Hinzu kommt, dass eine sichtbarere Positionierung eines Beitrags die Chance auf viele Seitenaufrufe erhöht – und Seitenaufrufe sind neben Verweildauer und Besucherzahl eine zentrale Metrik, anhand derer der Erfolg datenjournalistischer Projekte gemessen wird.

Zunehmend werden traditionelle Web-Aufmerksamkeitsmetriken wie Verweildauer durch andere quantitative Indikatoren aus Sozialen Medien ergänzt. Im Datenjournalismus sind Social-Web-Währungen wie „Shares“, „Likes“, „Fans“ oder „Follower“ doppelt relevant: einerseits bei der Veröffentlichung fertiger Projekte als Maßstab für deren Erfolg, andererseits bei Projekten, die sich inhaltlich mit dem Web beschäftigten, als Datenquelle für Analysen und Visualisierungen. Mit letzterem beschäftigt sich der folgende Abschnitt: mit datenjournalistischen Projekten, die ähnlichen Fragen wie kommunikationswissenschaftliche Internetforschung nachgehen.

#### 5.4 Stabilisierung digitaler datenjournalistischer Methoden

Automatisierungen, Fokus auf digitale Objekte und die Analysen von Web-Plattformen: Das sind einige zentrale Merkmale von *Digital Methods*, wie sie in Abschnitt zwei besprochen wurden. Diese Prinzipien sind auch vielen datenjournalistischen Projekten inhärent. So gibt es zahlreiche journalistische Analysen, die die Kommunikationsinhalte, das Verhalten oder die Netzwerke von NutzerInnen untersuchen. Das Themenspektrum ist dabei ähnlich breit wie in der kommuni-

kationswissenschaftlichen Internetforschung und reicht von Mappings von Nutzerinhalten bei Naturkatastrophen bis hin zur Analyse politischer und persönlicher Öffentlichkeiten. Aufgrund des relativ einfachen Datenzugangs greifen viele Analysen auf Social-Media-Dienste und hier insbesondere auf Twitter zurück.

Bei größeren Nachrichtenereignissen werden datenjournalistische Digital-Methods-Untersuchungen immer wieder herangezogen, um die Kommunikationsintensität oder Meinungen eines Teils der Bevölkerung zu erheben. Ein Datenjournalist (17) erklärt, warum er Social-Media-Daten für seine Berichterstattung heranzieht: „Jetzt haben wir die Möglichkeit, zumindest ein Gefühl dafür zu bekommen, wie tausende Menschen zu einem Thema denken.“ Beispielhaft für diese Herangehensweise ist das Projekt „Die Freunde von Pegida“ (Biermann et al., 2015), bei dem die Anhängerbasis der Facebook-Seite der ausländerfeindlichen Dresdner Organisation *Pegida* analysiert wurde. Nachdem der Ertrag klassischer soziologischer Erhebungen bislang „dünn“ sei, wandten sich die Autoren der Datenbasis der Fanseite zu, untersuchten etwa deren Wachstum, deren angegebene demografischen Besonderheiten und Vorlieben.<sup>14</sup> Dass diese Social-Media-Öffentlichkeiten nicht repräsentativ sind und mit ihrer Konstruktion auch Politik gemacht wird, ist den befragten ExpertInnen dabei durchaus bewusst. Die Interviews zeigten, dass insbesondere aufgrund des häufig vorherrschenden Ressourcenmangels Hoffnungen in Automatisierungen gesetzt werden: Einige der Interviewten würden gerne Scripts und Scraper so konfigurieren, dass sie manuelle Arbeit verringern, und die Veröffentlichung schneller bzw. annähernd in Echtzeit (etwa bei Social-Media-Daten) passieren kann.

Viele journalistische Digital-Methods-Projekte beschäftigen sich auch mit den persönlichen Öffentlichkeiten von Personen, die im Medieninteresse stehen. Exemplarisch seien die eingangs beschriebene Netzwerkanalyse des mutmaßlichen Boston-Marathon-Attentäters (Lotan, 2013) sowie die Analysen über die Twitteraktivität von Berliner (Tröger, 2012) und Schweizer (Thelitz et al., 2014) Abgeordneten angeführt. Ein gut dokumentiertes Beispiel für digitale Methoden im Datenjournalismus ist eine Untersuchung über die Verbreitung von Falschmeldungen auf Twitter während der Ausschreitungen in London im Sommer 2011 (Dant & Richards,

14 Diese Daten konnten ohne großen Aufwand per Schnittstelle von Facebook ausgelesen werden.

2011; Guardian Interactive Team et al., 2011). Für dieses Projekt arbeitete der Guardian mit drei britischen Universitäten zusammen. Insgesamt wurden 2,6 Millionen Tweets mit spezifischen Hashtags gesammelt, gefiltert und ein Teilkorpus davon manuell codiert. Mit Hilfe einer zeitbasierten interaktiven Visualisierung konnten die AutorInnen Momente aufzeigen, in denen Gerüchte auf Twitter verbreitet und entkräftet wurden (Procter et al., 2013).

## 6 Resümee

Dieser Aufsatz beleuchtete die Rolle von digitalen Methoden im Datenjournalismus, die mit Hilfe von Interviews untersucht wurden. In einem ersten Schritt wurden unter Rückgriff auf unter anderem Rogers (2013), Marres (2012), Latour (2005) und Schmidt (2012) die Grundzüge einer Methodologie formuliert, mithilfe derer die Anwendung von Digitalen Methoden im Journalismus untersucht werden sollte. Im Zentrum stand dabei die Frage, wie sich datenjournalistische Routinen manifestieren und welche Praktiken und heterogenen Akteure dabei vermitteln. Aufbauend auf qualitativen Analysen von elf semistrukturierten Interviews und Beobachtungsprotokollen aus Redaktionsbesuchen konnten einerseits neue Perspektiven auf Methoden und Routinen im Datenjournalismus gegeben werden, andererseits auch Arbeiten besprochen werden, in denen Daten aus dem Web den Ausgangspunkt für datenjournalistische Werke darstellen.

Datenjournalismus stellt sich als ein dynamisches Feld dar, das seine Routinen und Methoden in den letzten Jahren stark verändert und an die Digitalisierung angepasst hat. Was kann die digitale Kommunikationsforschung von den Entwicklungen des Datenjournalismus lernen?<sup>15</sup> Potenzial gibt es sowohl im Entwicklungsprozess als auch in der Ergebnispräsentation. Wie die Resultate dieser Arbeit darlegen, ist die enge Zusammenarbeit zwischen Personen mit unterschiedlichen Hintergründen (Journalismus, Design, Programmierung, etc.) zentral für das Gelingen eines Projekts. Die Einbeziehung sämtlicher Akteure von Beginn an, kurze Feedbackschleifen und die Anlehnung an Methoden der itera-

15 Mit der Frage, wie Journalismus von Digital Methods profitieren kann, beschäftigen sich Jonathan Gray und Liliana Bounegru (2014).

tiv-inkrementellen Software-Entwicklung ermöglichen es, im Datenjournalismus rasch zu präsentablen Ergebnissen zu kommen. Klar ist zugleich, dass in der Internetforschung in vielen Fällen andere Tempi notwendig sind, weil Wissenschaft andere Ansprüche an Validität stellt.

Im Bereich der Ergebnispräsentation liegt Potenzial für die Kommunikationswissenschaft in den interaktiven Darstellungsformen des Datenjournalismus: Viele Projekte zeichnen sich durch innovative interaktive Visualisierungen aus, über die Narrative vermittelt werden. In anderen Projekten wiederum können sich die RezipientInnen in den Ergebnissen wiederfinden und selbst einen Narrativ entwickeln. Auch die digitale Kommunikationsforschung kann Ähnliches in der Dissemination leisten oder Kooperationen mit Datenjournalismus-Teams in Redaktionen suchen. In den Vergangenheit gab es einige erfolgreiche Kollaborationsprojekte, von denen sowohl Journalismus als auch Wissenschaft profitiert haben. Ein Beispiel ist das oben diskutierte „Reading the Riots“-Projekt des *Guardian*.

Hinsichtlich der Methodentransparenz und der Offenlegung von Algorithmen kann Datenjournalismus durchaus Vorbild für die Kommunikationswissenschaft sein. Profitieren könnte sie hier darüber hinaus auch von den Digital Humanities und der Open-Science-Bewegung, die sich schon viele Jahre mit diesen Themen auseinandersetzen (Ausserhofer, 2015).

Auch wäre es sinnvoll, wenn der Datenjournalismus die Standardisierung seiner digitalen Methoden weiter vorantreibt. Langsam kristallisieren sich etwa neue Erzählformen heraus. Eine Vereinheitlichung von digitalen Methoden und Technologien ist wünschenswert, um die längerfristige Zugänglichkeit zu gewährleisten und Archivierung zu ermöglichen.

*Mag. (FH) Julian Ausserhofer* ist wissenschaftlicher Mitarbeiter am Institut für Journalismus und Public Relations der FH Joanneum Graz und Doktorand am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

## Quellenverzeichnis

Anderson, C. W. (2013). Towards a Sociology of Computational and Algorithmic Journalism. *New Media & Society*, 15(7), 1005-1021. doi: 10.1177/1461444812465137

- Appelgren, E., & Nygren, G. (2014). Data Journalism in Sweden: Introducing New Methods and Genres of Journalism into "Old" Organizations. *Digital Journalism*, 2(3), 394-405. doi: 10.1080/21670811.2014.884344
- Ausserhofer, J. (2015). Die Datenbank verdient die Hauptrolle: Aspekte einer Methodologie für Open Digital Humanities. In S. Eichhorn, B. Oberreither, M. Rauchenbacher, I. Schwentner, & K. Serles (Hrsg.), *Aufgehoben? Speicherorte, -diskurse und -medien*. Würzburg: Königshausen & Neumann.
- Berg, B. L. (2001). *Qualitative Research Methods for the Social Sciences*. Boston: Allyn and Bacon.
- Berry, D. M. (2011). The Computational Turn: Thinking About the Digital Humanities. *Culture Machine*, 12, 1-22. Abgerufen von <http://www.culturemachine.net/index.php/cm/article/viewArticle/440>
- Biermann, K., Blicke, P., & Venohr, S. (2015, 5. Februar). Dürfen wir vorstellen: Die Freunde von Pegida. *Die Zeit*. Abgerufen von <http://www.zeit.de/gesellschaft/zeitgeschehen/2015-02/wer-ist-pegida-facebook-daten/>
- Bodle, R. (2011). Regimes of Sharing: Open APIs, Interoperability, and Facebook. *Information, Communication & Society*, 14(3), 320-337. doi: 10.1080/1369118X.2010.542825
- Borger, M., van Hoof, A., Costera Meijer, I., & Sanders, J. (2013). Constructing Participatory Journalism as a Scholarly Object. *Digital Journalism*, 1(1), 117-134. doi: 10.1080/21670811.2012.740267
- Bounegru, L., & Gray, J. (2014, September). Mapping Issues with the Web: An Introduction to Digital Methods. Abgerufen von <http://www.slideshare.net/jwyg/mapping-issues-with-the-web-an-introduction-to-digital-methods>
- Bucher, T. (2013). Objects of Intense Feeling: The Case of the Twitter API. *Computational Culture*, (3). Abgerufen von <http://computationalculture.net/article/objects-of-intense-feeling-the-case-of-the-twitter-api>
- Bradshaw, P. (2011a, 7. Juli). The inverted pyramid of data journalism. Abgerufen von <http://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>
- Bradshaw, P. (2011b, 13. Juli). 6 ways of communicating data journalism (The inverted pyramid of data journalism part 2). Abgerufen von <http://onlinejournalismblog.com/2011/07/13/the-inverted-pyramid-of-data-journalism-part-2-6-ways-of-communicating-data-journalism/>

- Coddington, M. (2014). Clarifying Journalism's Quantitative Turn: A Typology for Evaluating Data Journalism, Computational Journalism, and Computer-Assisted Reporting. *Digital Journalism*, Online First. doi: 10.1080/21670811.2014.976400
- Cohen, S., Hamilton, J. T., & Turner, F. (2011). Computational Journalism. *Communications of the ACM*, 54(10), 66-71. doi: 10.1145/2001269.2001288
- Dant, A., & Richards, J. (2011, 8. Dezember). Behind the Rumours: How We Built Our Twitter Riots Interactive. *The Guardian*. Abgerufen von <http://www.guardian.co.uk/news/datablog/2011/dec/08/twitter-riots-interactive>
- De Maeyer, J., Libert, M., Domingo, D., Heinderyckx, F., & Cam, F. L. (2014). Waiting for Data Journalism: A Qualitative Assessment of the Anecdotal Take-Up of Data Journalism in French-Speaking Belgium. *Digital Journalism*, Online First. doi: 10.1080/21670811.2014.976415
- Domingo, D., & Paterson, C. (Hrsg.) (2011). *Making Online News Volume 2: Newsroom Ethnographies in the Second Decade of Internet Journalism*. New York: Peter Lang.
- Flew, T., Daniel, A., & Spurgeon, C. L. (2010). The Promise of Computational Journalism. In K. McCallum (Hrsg.), *Media, Democracy and Change: Refereed Proceedings of the Australian and New Zealand Communication Association Annual Conference* (S. 1-19). Canberra. Abgerufen von <http://eprints.qut.edu.au/39649/>
- Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Fink, K., & Anderson, C. W. (2014). Data Journalism in the United States: Beyond the "usual suspects." *Journalism Studies*, Online First. doi: 10.1080/1461670X.2014.939852
- Guardian Interactive Team, Procter, R., Vis, F., & Voss, A. (2011, 7. Dezember). How Riot Rumours Spread on Twitter. *The Guardian*. Abgerufen von <http://www.theguardian.com/uk/interactive/2011/dec/07/london-riots-twitter>
- Gynnild, A. (2014). Journalism Innovation Leads to Innovation Journalism: The Impact of Computational Exploration on Changing Mindsets. *Journalism*, 15(6), 713-730. doi: 10.1177/1464884913486393
- Haas, H. (1999). *Empirischer Journalismus: Verfahren zur Erkundung gesellschaftlicher Wirklichkeit*. Wien: Böhlau.
- Hamilton, J. T., & Turner, F. (2009). Accountability Through Algorithm: Developing the Field of Computational Journalism. Presented at the Behavioral Sciences Summer Workshop. Abgerufen von <http://dewitt.sanford.duke.edu/wp-content/uploads/2011/12/About-3-Research-B-cj-1-finalreport.pdf>

- Howard, A. B. (2014). *The Art and Science of Data-Driven Journalism*. Tow Center for Digital Journalism. Abgerufen von <http://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Data-Driven-Journalism.pdf>
- Karlsen, J., & Stavelin, E. (2014). Computational Journalism in Norwegian Newsrooms. *Journalism Practice*, 8(1), 34-48. doi: 10.1080/17512786.2013.813190
- Latour, B. (1988). *The Pasteurization of France*. (A. Sheridan & J. Law, Übers.). Cambridge: Harvard University Press.
- Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York: Oxford University Press.
- Law, J. (2004). *After Method: Mess in Social Science Research*. London: Routledge.
- Lewis, S. C., & Usher, N. (2014). Code, Collaboration, and the Future of Journalism: A Case Study of the Hacks/Hackers Global Network. *Digital Journalism*, 2(3), 383-393. doi: 10.1080/21670811.2014.895504
- Lotan, G. (2013, 25. April). When Your Twitter Friend Turns Out To Be The Boston Bomber. *Digg*. Abgerufen von <http://digg.com/originals/dzhokhar-tsarnaev-twitter-map>
- Lorenz, M. (2010). Status and Outlook for Data-Driven Journalism. In European Journalism Centre (Hrg.), *Data-Driven Journalism: What Is There to Learn?* (S. 8-17). Amsterdam.
- Madrigal, A. C. (2013, 19. April). #BostonBombing: The Anatomy of a Misinformation Disaster. *The Atlantic*. Abgerufen von <http://www.theatlantic.com/technology/archive/2013/04/-bostonbombing-the-anatomy-of-a-misinformation-disaster/275155/>
- Manovich, L. (2013). *Software Takes Command*. New York: Bloomsbury.
- Marres, N. (2012). The Redistribution of Methods: On Intervention in Digital Social Research, Broadly Conceived. *The Sociological Review*, 60(S1), 139-165. doi: 10.1111/j.1467-954X.2012.02121.x
- Meyer, P. (2002/1973). *Precision Journalism: A Reporter's Introduction to Social Science Methods* (4. Auflage). Oxford: Rowman & Littlefield.
- Münch, R., & Schmidt, J. (2005). Medien und sozialer Wandel. In M. Jäckel (Hrsg.), *Mediensoziologie: Grundfragen und Forschungsfelder* (S. 201-218). Wiesbaden: VS Verlag.
- Niederer, S., & van Dijck, J. (2010). Wisdom of the Crowd or Technicity of Content? Wikipedia as a Sociotechnical System. *New Media & Society*, 12(8), 1368-1387. doi: 10.1177/1461444810365297



- Niederer, S. (2013). "Global Warming Is Not a Crisis!": Studying Climate Change Skepticism on the Web. *NECSUS. European Journal of Media Studies*, 2(1), 83-112. doi: 10.5117/NECSUS2013.1.NIED
- Parasie, S., & Dagiral, E. (2013). Data-Driven Journalism and the Public Good: "Computer-Assisted-Reporters" and "Programmer-Journalists" in Chicago. *New Media & Society*, 15(6), 853-871. doi: 10.1177/1461444812463345
- Preston, J., & Roston, M. (2013, 20. April). A Closer Look at the Bombing Suspect's Twitter Account. *New York Times*. Abgerufen von <http://thelede.blogs.nytimes.com/2013/04/20/dzhokhar-tsarnaevs-jahar-twitter-account-prompts-scrutiny/>
- Procter, R., Vis, F., & Voss, A. (2013). Reading the Riots on Twitter: Methodological Innovation for the Analysis of Big Data. *International Journal of Social Research Methodology*, 16(3), 197-214. doi: 10.1080/13645579.2013.774172
- Rieder, B., & Röhle, T. (2012). Digital Methods: Five Challenges. In D. M. Berry (Hrsg.), *Understanding Digital Humanities* (S. 67-84). Basingstoke: Palgrave Macmillan.
- Rogers, R. (2009). *The End of the Virtual: Digital Methods*. Amsterdam: Vossiuspers UvA.
- Rogers, R. (2010). Internet Research: The Question of Method – A Keynote Address from the YouTube and the 2008 Election Cycle in the United States Conference. *Journal of Information Technology & Politics*, 7, 241-260. doi: 10.1080/19331681003753438
- Rogers, R. (2013). *Digital Methods*. Cambridge: MIT Press.
- Rogers, R. (2014). Political Research in the Digital Age. *International Public Policy Review*, 8(1), 73-87.
- Rogers, S. (2011, 7. April). Data Journalism Broken down: What We Do to the Data Before You See It. Abgerufen von <http://www.theguardian.com/news/datablog/2011/apr/07/data-journalism-workflow>
- Royal, C. (2012). The Journalist as Programmer: A Case Study of the New York Times Interactive News Technology Department. *#ISOJ The Official Research Journal of the International Symposium on Online Journalism*, 2(1). Abgerufen von <https://online.journalism.utexas.edu/2010/papers/Royal10.pdf>
- Rühl, M. (1981). Journalismus und Wissenschaft: Anmerkungen zu ihrem Wirklichkeitsverständnis. *Rundfunk und Fernsehen*, 2(3), 211-222.

- Schmidt, R. (2012). *Soziologie der Praktiken: Konzeptionelle Studien und empirische Analysen*. Berlin: Suhrkamp.
- Seward, Z. M., Mirani, L., & King, R. (2013, 20. April). We know when Dzhokhar Tsarnaev sleeps. *Quartz*. Abgerufen von <http://qz.com/76442/we-know-when-dzhokhar-tsarnaev-sleeps/>
- Strauss, A., & Corbin, J. M. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks: Sage.
- Thelitz, N., Kohli, A., Fichter, A., Kobler, S., Preusse, T., Schudel, B., ... Interactive Things. (2014, 8. September). Ungleiches Gezwitscher aus dem Parlament. *Neue Zürcher Zeitung*. Abgerufen von <http://www.nzz.ch/schweiz/ungleiches-gezwitscher-aus-dem-parlament-1.18378048>
- Tröger, J. (2012, 11. Oktober). Wie die Berliner Abgeordneten Twitter nutzen. *Berliner Morgenpost*. Abgerufen von <http://www.morgenpost.de/berlin-aktuell/article109759023/Wie-die-Berliner-Abgeordneten-Twitter-nutzen.html>
- Weber, W., & Rall, H. (2013). "We Are Journalists": Production Practices, Attitudes and a Case Study of the New York Times Newsroom. In W. Weber, M. Burmester, & R. Tille (Hrsg.), *Interaktive Infografiken* (S. 161-172). Berlin: Springer. doi: 10.1007/978-3-642-15453-9\_9
- Weinacht, S., & Spiller, R. (2014). Datenjournalismus in Deutschland: Eine explorative Untersuchung zu Rollenbildern von Datenjournalisten. *Publizistik*, 59(4), 411-433. doi: 10.1007/s11616-014-0213-5
- Weinstein, M. (2006). TAMS Analyzer Anthropology as Cultural Critique in a Digital Age. *Social Science Computer Review*, 24(1), 68-77. doi: 10.1177/0894439305281496
- Weltevrede, E., & Helmond, A. (2012). Where Do Bloggers Blog? Platform Transitions Within the Historical Dutch Blogosphere. *First Monday*, 17(2). doi: 10.5210/fm.v17i2.3775
- Young, M. L., & Hermida, A. (2014). From Mr. and Mrs. Outlier to Central Tendencies: Computational Journalism and Crime Reporting at the Los Angeles Times. *Digital Journalism*, Online First. doi: 10.1080/21670811.2014.976409

## **Danksagung**

Die Forschung, die dieser Arbeit zugrunde liegt, wurde ermöglicht durch Mittel der Universität Wien, der FH Joanneum sowie der Österreichischen Forschungsförderungsgesellschaft (FFG) (Projekt „Visual Analytics in Data-driven Journalism“ (VALiD), Projektnummer: 845598, Fördergeber: Bundesministerium für Verkehr, Innovation und Technologie). Bedanken möchte ich mich bei meinen InterviewpartnerInnen sowie Wolfgang Aigner, Robert Gutounig, Axel Maireder, Sarah Matiassek, Alice Miller, Wiebke Loosen, Alexander Rind, Elena Rudowsky und Christina Schumann. Dank gebührt darüber hinaus dem Alexander von Humboldt Institut für Internet und Gesellschaft in Berlin sowie dem Center for Science, Technology, Medicine and Society an der University of California, Berkeley, wo Teile der Forschung und Textproduktion stattgefunden haben.